

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/501, 183
Source: PCT
Date Processed by STIC: 11/9/04

ENTERED

CRF Errors Edited by the STIC Systems Branch

Serial Number: 10/506/183

CRF Edit Date: 11/9/04
Edited by: TC

☒ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

☐ Corrected the SEQ ID NO. Sequence numbers edited were:

☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

☒ Deleted: ☒ Invalid beginning/end-of-file text ; _____ page numbers

☐ Inserted mandatory headings/numeric identifiers, specifically:

☐ Moved responses to same line as heading/numeric identifier, specifically:

☐ Other:



PCT

RAW SEQUENCE LISTING

DATE: 11/09/2004

PATENT APPLICATION: US/10/501,183

TIME: 17:50:38

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\11092004\J501183.raw

3 <110> APPLICANT: Karatzas, Costas
 4 Turcotte, Carl
 6 <120> TITLE OF INVENTION: METHODS OF PRODUCING SILK POLYPEPTIDES AND PRODUCTS THEREOF
 8 <130> FILE REFERENCE: 9529-008-999
 10 <140> CURRENT APPLICATION NUMBER: 10/501,183
 C--> 11 <141> CURRENT FILING DATE: 2004-07-09
 13 <150> PRIOR APPLICATION NUMBER: US 60/347,509
 14 <151> PRIOR FILING DATE: 2002-01-11
 16 <160> NUMBER OF SEQ ID NOS: 48
 18 <170> SOFTWARE: PatentIn version 3.2
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 646
 22 <212> TYPE: PRT
 23 <213> ORGANISM: Artificial sequence
 25 <220> FEATURE:
 26 <223> OTHER INFORMATION: MaSpI polypeptide
 28 <400> SEQUENCE: 1
 30 Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln
 31 1 5 10 15
 34 Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln
 35 20 25 30
 38 Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly
 39 35 40 45
 42 Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly
 43 50 55 60
 46 Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Val Gly Gln
 47 65 70 75 80
 50 Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala
 51 85 90 95
 54 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser
 55 100 105 110
 58 Gln Gly Ala Gly Arg Gly Gly Ser Gly Gly Gln Gly Ala Gly Ala Ala
 59 115 120 125
 62 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly
 63 130 135 140
 66 Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala
 67 145 150 155 160
 70 Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly
 71 165 170 175
 74 Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser
 75 180 185 190
 78 Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala
 79 195 200 205

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,183

DATE: 11/09/2004

TIME: 17:50:38

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\11092004\J501183.raw

```

82 Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln
83      210                215                220
86 Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala
87 225                230                235                240
90 Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly
91      245                250                255
94 Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Val Gly Ala Gly Gln
95      260                265                270
98 Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu
99      275                280                285
102 Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly
103      290                295                300
106 Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly
107 305                310                315                320
110 Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly
111      325                330                335
114 Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Asn Gln Gly Ala Gly
115      340                345                350
118 Arg Gly Gly Gln Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln
119      355                360                365
122 Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu
123      370                375                380
126 Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly
127 385                390                395                400
130 Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly
131      405                410                415
134 Tyr Gly Gly Leu Gly Ser Gln Gly Ser Gly Arg Gly Gly Leu Gly Gly
135      420                425                430
138 Gln Gly Ala Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly
139      435                440                445
142 Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala
143      450                455                460
146 Ala Ala Gly Gly Val Arg Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln
147 465                470                475                480
150 Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala
151      485                490                495
154 Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Val
155      500                505                510
158 Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Gly
159      515                520                525
162 Gly Ala Gly Gln Gly Gly Tyr Gly Gly Val Gly Ser Gly Ala Ser Ala
163      530                535                540
166 Ala Ser Ala Ala Ala Ser Arg Leu Ser Ser Pro Gln Ala Ser Ser Arg
167 545                550                555                560
170 Val Ser Ser Ala Val Ser Asn Leu Val Ala Ser Gly Pro Thr Asn Ser
171      565                570                575
174 Ala Ala Leu Ser Ser Thr Ile Ser Asn Val Val Ser Gln Ile Gly Ala
175      580                585                590
178 Ser Asn Pro Gly Leu Ser Gly Cys Asp Cys Leu Ile Gln Ala Leu Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,183

DATE: 11/09/2004

TIME: 17:50:38

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\11092004\J501183.raw

```

179          595          600          605
182 Glu Val Val Ser Ala Leu Ile Gln Ile Leu Gly Ser Ser Ser Ile Gly
183          610          615          620
186 Gln Cys Asn Tyr Gly Ser Ala Gly Gln Ala Thr Gln Ile Val Gly Gln
187 625          630          635          640
190 Ser Val Tyr Gln Ala Leu
191          645
194 <210> SEQ ID NO: 2
195 <211> LENGTH: 627
196 <212> TYPE: PRT
197 <213> ORGANISM: Artificial sequence
199 <220> FEATURE:
200 <223> OTHER INFORMATION: MaSpII polypeptide
202 <400> SEQUENCE: 2
204 Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro
205 1          5          10          15
208 Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala
209          20          25          30
212 Ala Ala Ala Ala Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro
213          35          40          45
216 Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly
217 50          55          60
220 Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala Gly
221 65          70          75          80
224 Ser Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro
225          85          90          95
228 Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala
229          100          105          110
232 Ala Ala Ala Ser Ala Ala Ala Ser Ala Glu Ser Gly Gln Gln Gly Pro
233          115          120          125
236 Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly
237 130          135          140
240 Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly
241 145          150          155          160
244 Pro Gly Ser Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gln
245          165          170          175
248 Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr
249          180          185          190
252 Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala
253          195          200          205
256 Ala Ala Ala Ala Ser Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly
257 210          215          220
260 Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Leu
261 225          230          235          240
264 Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gln
265          245          250          255
268 Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro
269          260          265          270
272 Gly Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gly Tyr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/501,183

DATE: 11/09/2004

TIME: 17:50:38

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\11092004\J501183.raw

```

273          275          280          285
276 Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly
277          290          295          300
280 Pro Ser Gly Ala Gly Ser Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly
281 305          310          315          320
284 Gln Gln Gly Leu Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly
285          325          330          335
288 Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Ser Ala
289          340          345          350
292 Ser Ala Ala Ala Ala Ala Ala Gly Pro Gly Gln Gln Gly Pro Gly Gly
293          355          360          365
296 Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ser Ala
297          370          375          380
300 Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gly Tyr Gly Pro Gly Gln
301 385          390          395          400
304 Gln Gly Pro Gly Gly Tyr Ala Pro Gly Gln Gln Gly Pro Ser Gly Pro
305          405          410          415
308 Gly Ser Ala Ser Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gly
309          420          425          430
312 Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Ala Pro Gly Gln Gln
313          435          440          445
316 Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala
317          450          455          460
320 Gly Pro Gly Gly Tyr Gly Pro Ala Gln Gln Gly Pro Ser Gly Pro Gly
321 465          470          475          480
324 Ile Ala Ala Ser Ala Ala Ser Ala Gly Pro Gly Gly Tyr Gly Pro Ala
325          485          490          495
328 Gln Gln Gly Pro Ala Gly Tyr Gly Pro Gly Ser Ala Val Ala Ala Ser
329          500          505          510
332 Ala Gly Ala Gly Ser Ala Gly Tyr Gly Pro Gly Ser Gln Ala Ser Ala
333          515          520          525
336 Ala Ala Ser Arg Leu Ala Ser Pro Asp Ser Gly Ala Arg Val Ala Ser
337          530          535          540
340 Ala Val Ser Asn Leu Val Ser Ser Gly Pro Thr Ser Ser Ala Ala Leu
341 545          550          555          560
344 Ser Ser Val Ile Ser Asn Ala Val Ser Gln Ile Gly Ala Ser Asn Pro
345          565          570          575
348 Gly Leu Ser Gly Cys Asp Val Leu Ile Gln Ala Leu Leu Glu Ile Val
349          580          585          590
352 Ser Ala Cys Val Thr Ile Leu Ser Ser Ser Ile Gly Gln Val Asn
353          595          600          605
356 Tyr Gly Ala Ala Ser Gln Phe Ala Gln Val Val Gly Gln Ser Val Leu
357          610          615          620
360 Ser Ala Phe
361 625
364 <210> SEQ ID NO: 3
365 <211> LENGTH: 625
366 <212> TYPE: PRT
367 <213> ORGANISM: Artificial sequence

```

RAW SEQUENCE LISTING

DATE: 11/09/2004

PATENT APPLICATION: US/10/501,183

TIME: 17:50:38

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\11092004\J501183.raw

369 <220> FEATURE:

370 <223> OTHER INFORMATION: ADF-3 polypeptide

372 <400> SEQUENCE: 3

```

374 Gly Ser Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly
375 1          5          10          15
378 Pro Gly Gln Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala
379          20          25          30
382 Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro Ser Gln
383          35          40          45
386 Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Arg Tyr Gly Pro
387          50          55          60
390 Gly Ala Ser Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser
391 65          70          75          80
394 Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ser Ser
395          85          90          95
398 Ala Ala Ala Ala Ala Ala Gly Gly Asn Gly Pro Gly Ser Gly Gln Gln
399          100          105          110
402 Gly Ala Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Ala Ser Ala
403          115          120          125
406 Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly
407          130          135          140
410 Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala
411 145          150          155          160
414 Ser Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln
415          165          170          175
418 Gly Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly
419          180          185          190
422 Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly
423          195          200          205
426 Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly
427          210          215          220
430 Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly
431 225          230          235          240
434 Gly Tyr Gly Pro Gly Tyr Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro
435          245          250          255
438 Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ser Ala Ala
439          260          265          270
442 Ser Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro Gly Gln Gln
443          275          280          285
446 Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala
447          290          295          300
450 Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro Gly
451 305          310          315          320
454 Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly
455          325          330          335
458 Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly
459          340          345          350
462 Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro
463          355          360          365

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/501,183

DATE: 11/09/2004

TIME: 17:50:39

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\11092004\J501183.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date